AMEA-nın Xəbərləri (biologiya və tibb elmləri), cild 68, №2, səh. 23-28 (2013)

Functional Genomics And Forward And Reverse Genetics Approaches For Identification Of Important QTLs In Plants

S.M. Zakhrabekova[a], S. Gough[a], L.Lundh[b], M.Hansson[a]

[a]*Carlsberg Laboratory, 1799 Copenhagen V, Denmark*
[b]*Department of Cell and Molecular Biology, Gothenburg University, Box 462, 405 30 Gothenburg, Sweden; E-mail: zakhrabekova@yahoo.se*

The main goal of functional genomics is to understand how the organism's genotype is expressed as its phenotype. Since mutations can interfere with cellular processes, induced mutations are key probes for understanding gene function.

Key words: *functional genomics, forward genetics, reverse genetics, QTL*

Forward and reverse genetics approaches are used to determine the function of genes.

While forward genetics refers to the identification and characterization of the gene that is responsible for the mutant phenotype, the goal of reverse genetics is to examine the effect of induced mutationor altered expression of a particular gene and to understand the gene function (Ahringer, 2006) (Fig. 1).

The aim of this article is to review the methods and approaches of reverse and forward genetics, as well as to review online tools and progress in this field of genetic studies.

Reverse Genetics

Gene silencing and homologous recombination are two commonly used approaches used for targeted gene mutation, in contrast to non-targeted disruptions of genes achieved by transposon mediated and chemical mutagenesis. (fig.2) For such a model plant as *Arabidopsis,* T-DNA insertion mutants have been produced and are available for researchers(Krysan et al., 1999). It should be noted that transpositions are not completely random(Krysan et al., 2002) and thus mutation of all genes is not guaranteed. However the mutants are a valuable research tools for understanding the function of the gene. The required T-DNA insertion mutant can usually be ordered and detailed phenotypic analysis performed.

Below we present an overview of two novel approaches of reverse genetics: 1) targeted gene silencing by RNA interference and 2) TILLING (Targeting Induced Local Lesions in Genomes) - a recently developed reverse genetics technique.

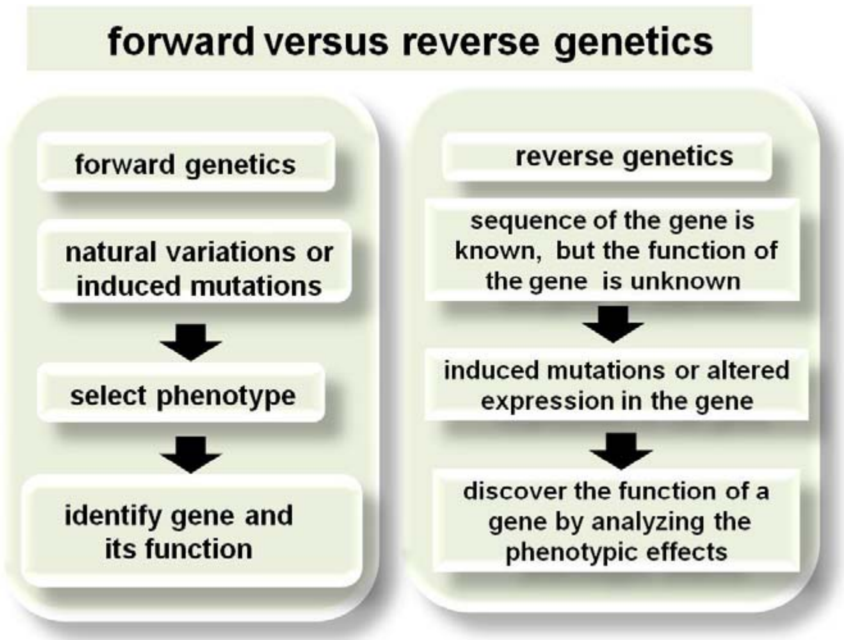


Figure1. The difference between forward and reverse genetic techniques is the starting point: the phenotype or the gene

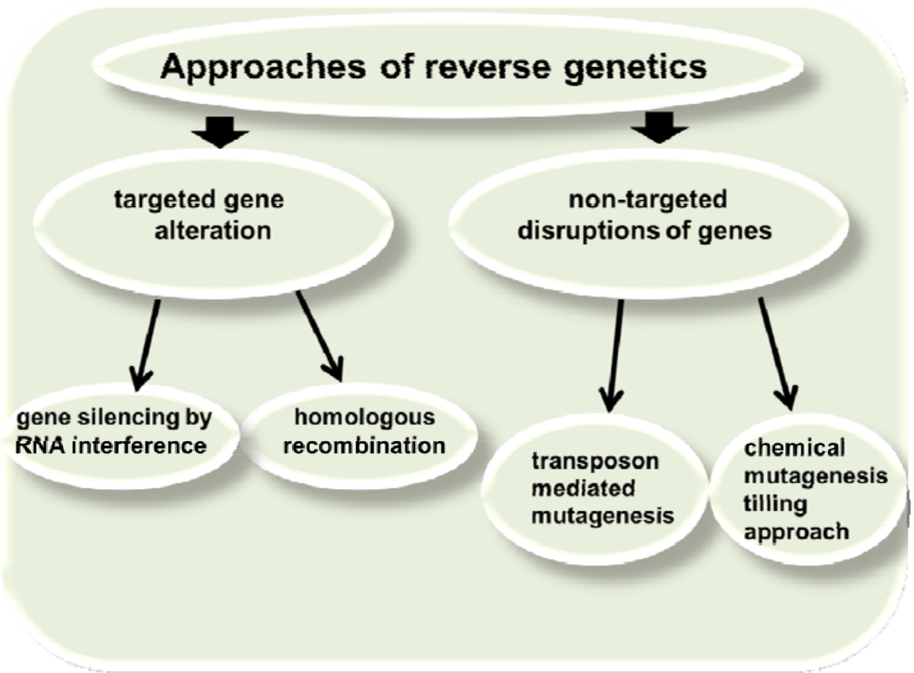


Figure 2. Approaches of reverse genetics

Targeted gene silencing

Gene silencing by RNA interference (RNAi) is one of the most exciting breakthrough of the past decade in functional genomics and promises to be a very useful instrument for therapeutic gene silencing.

The phenomenon of RNAi was first discovered during experiments associated with changes in pigmentation in the petunia plants. Introducing extra copies of a pigment biosynthesis gene did not increase the color intensity in the flower as was expected, but the flowers became less colorful than the wild flowers (de Lange et al., 1995; Hannon, 2002; Bushman, 2003; Foubister, 2003). The term RNAi was first introduced after Andrew Fire and Craig C. Mello discovered that injection of dsRNA into the nematode worm *Caenorhabditis elegans*cause the specific silencing of genes highly homologous to the supplied sequence (Fire et al., 1998; Elbashir et al., 2001).

Double-stranded RNA (dsRNA) triggers the RNAi process and can be endogenous or exogenously introduced into the cells(Shuey et al., 2002).The basis of the RNAi process, production of the functionally similar endogenously produced siRNAs, is quite similar in many organisms and the enzymes required for this process show high inter species homology (Hamilton and Baulcombe, 1999; Paddison et al., 2002). Processing of dsRNA precursors into small interfering RNAs (siRNAs) is mediated by special dsRNA-specific RNase-III-type endonucleases, known as Dicer. This results in formation of 21-25 nucleotide double stranded RNA duplexes with symmetric 2-3 nucleotide 3' overhangs, which are called small interfering RN-siRNA.The siRNAs are afterwards incorporated into the RNA-induced silencing complex (RISC), where an RNA helicase unwinds the inactive double-stranded siRNA, converting it to an active single-stranded form (Nykanen et al., 2001; Hannon, 2002; Plasterk, 2002). Nevertheless only one strand, known as the guide strand is stabilized in RISC complex, while the passenger siRNA strand is degraded. (Gregory RI., et al., 2005). An active RISC complex uses the guide siRNA to find and destroy the complementary sequence of mRNA, causing in turn gene silencing (Bushman F, 2003: 49, Nykänen A., et al., 2001).

In plants in contrast to other organisms miRNA have perfect or near perfect complementarity to their targets(Axtell et al., 2011). Thus plant siRNAs are easily designed**.**

TILLING (Targeting Induced Local Lesions in Genomes)

Generation of mutated lines by transposons, T-DNA or RNA interference is technically difficult in some organisms.The difficulty comes due to the lack of an efficient transformation system and due to the large genome for some organisms like barley (Bennett and Smith, 1976; Ahringer, 2006; Chawade et al., 2010). One way to increase

variation in the breeding process would be to use radiation or chemical mutagens such as EMS (ethyl methanesulfonate). The mutagenic substance EMS preferentially alkylates guanine bases. The resulting O-6-ethyl guanine paired with cytosine is misread by the DNA-replicating polymerase which insertsa thymine residue instead of a cytosine residue. This results in G-C base-pairs (bps) being mutated to A-T (Hoffmann, 1980). Mutations in coding regions can be silent, missense or nonsense and mutations in non-coding promoter or intron regions can resultin up- or down-regulation of transcription (Rose and Beliakoff, 2000). TILLING (Targeting Induced Local Lesions in Genomes) is a recently developed reverse genetics technique, based on the use of a mismatch-specific endonuclease (*CelI*), which finds mutations in a target gene containing aheteroduplex formation (Henikoff et al., 2004; Gilchrist et al., 2006; Chawade et al., 2010). If the mutation frequency is high and the population size large enough, mutated alleles of most, if not all, genes will be present in the population. The technique involves PCR amplification of the target gene using fluorescently labeled primers, formation of DNA heteroduplex between wild type and mutant alleles (PCR products, corresponding to the mutant and wild type alleles are heated and then slowly cooled), followed by endonuclease digestion specifically cleaving at the site of an EMS induced mismatch. The sizes of the amplicon cleavage fragments are often analyzed by a Li-COR (McCallum CM, 2000) or MALDI-TOF (Matrix-assisted laser desorption/ionization-time-of-flight mass spectrometer) (Chawade et al., 2010) system. It is possible to apply TILLING to genetically complicated crops, such as wheat for example (Slade et al., 2005).

One of the greatest benefits of the TILLING approach is that it does not involve genetic manipulations, that results in Genetically Modified Organisms (GMO), which are not legal for agricultural applications in many countries.

Forward genetics

The aim of forward genetics is to determine the genetic basis of observed phenotypic variation. To generate random mutations in an organism, various approaches are exploited for example X-rays, ultraviolet irradiation and chemical treatment. These gene disruptions are followed by selection of aberrant phenotypes, associated with various traits, such as high-yield, early maturity, lodging resistance, disease resistance, drought tolerance, cold tolerance, toxic metal resistance, etc. After mutants are identified, they need to be classified. The aim is to gather mutants into complementation groups by using allelism tests. Such groups of multiple independent mutant alleles can efficiently be used to validate a candidate gene.

One example of such collection is the Scandinavian barley mutant collection. The generation of this collection has started in 1928 by the Swedish geneticists Hermann Nilsson-Ehle and Åke Gustafsson. In the mid 1930-ies, the first viable mutations were observed and notable among them arehigh-yielding, early maturity, dense spike, tillering capacity, straw-stiffness, seed-size and mutants useful for understanding basic agronomically important traits such as photosynthetic capacity and protective outer barrier formation(Lundqvist, 2005).In this way the barley mutants became very important for breeding-improved varieties and for subsequent genetic studies.

Genetic mapping

The goal of genetic mapping is to identify the locus of the gene responsible for the trait of interest. The first step in all mapping studies is to find markers that are linked with the trait. Physical linkage will lead to co-inheritance of markers, while recombination events will break these associations. The next steps are to develop appropriate mapping populations; screen parents for marker polymorphism and genotype mapping population. Afterwards a linkage analysis is performed to find out recombination frequencies between markers which in turn lead to the fine mapping of the location of the gene of interest.

If the genome of the plant of interest is not fully sequenced, the synteny between physical and genetic maps of closely related plants, with sequenced genome enables the assessment of the gene content at the fine mapped locus. The following databases and their online genome browsers and blast search capabilities are essential for these syntenic studies:

- NCBI:http://www.ncbi.nlm.nih.gov
- Phytozome: comparative genomics of plants. http://www.phytozome.net/
- PlantGDB: plant genome database: http://mips.helmholtz-muenchen.de/plant/genomes.jsp

Expression analyses with microarray

DNA microarray is one of the most efficient methods for gene expression analysis (Gregory et al., 2008; Morohashi et al., 2009;Park et al., 2004; Petersen et al., 2005; Schena et al., 1995;Zhu et al., 2012). It was further shown that microarray is a very promising technology for identification of

genes in transcription deficient mutants (Zakhrabekova S, 2002; Zakhrabekova et al., 2007). The approach of using phenotypically similar mutants minimizes the number of candidate genes for sequencing, due to the reduction of genes which are secondarily affected by the mutation. Both cDNA and Affymetrix microarray platforms are able to successfully pinpoint the gene which is down- or up-regulated due to induced or naturally occurred mutation events.

It was also shown that nonsense-mediated mRNA decay in barley mutants expands the number of mutants that can be used for gene identification by the microarray approach(Gadjieva et al., 2004).

Useful data base for gene expression studies;

- PLEXdb (Plant Expression Database) is a unified gene expression resource for plants and plant pathogens: http://www.plexdb.org/
- Planet is a gene expression database for the plants: barley, rice, wheat, Arabidopsis, Medicago, poplar and soybean based on Affymetrix contigs: http://aranet.mpimp-golm.mpg.de/

Different microarrays technologies:

- Affymetrix: http://www.affymetrix.com/estore/
- Agilent: http://www.genomics.agilent.com/GenericB.aspx?PageType=Custom&SubPageType=Custom&PageID=2011

Candidate gene approach

This approach is appropriate for plants where mutant collections, represented by multiple independent mutant alleles are available.

The major difficulty with this approach is that in order to choose a potential candidate gene for the mutation, researchers must already have an understanding of the mechanisms underlying the phenotypic disorder. Very good "educated guesses" can be done if a study of similar mutants has been performed in another related plant and the corresponding orthologous gene has been identified. Then this gene can be a potential candidate for the mutation in the investigated plant and the principle proof that this candidate gene is responsible for the observed phenotype is coming from comparative sequence analysis of all available mutant alleles in the particular locus (Zakhrabekova et al., 2012).

Alternative methods which can be used to hunt a gene responsible for a mutant phenotype.

The RNA-seq method is also called "Whole Transcriptome Shotgun Sequencing" ("WTSS") (Morin et al., 2008). This is a high-throughput sequencing technologies to sequence cDNA in order to get information about the RNA content in the cells. Since converting RNA into cDNA by using reverse transcriptase might introduce mutations, single-molecule direct RNA sequencing technology has been developed. The sequences of all RNA in the mutant are then compared with the wild-type indicate mutant candidates. This method also requires a number of different mutant alleles to give a reliable answer (Ozsolak and Milos, 2011).

Exome sequencing is a powerful method to selectively sequence the coding regions of the genome as a less costly alternative to whole genome sequencing (Ng et al. 2009). This method can be combined with target-enrichment strategies, which give possibility to selectively capture genomic regions of interest from a DNA sample prior to sequencing (Basiardes et al., 2005).

Identification of mutations by this method requires as well as in RNA-seq a number of different mutant alleles to give a trust worthy answer.

REFERENCES

Ahringer J. (2006) Reverse genetics. In: Wormbook (Editor The C. elegans Research Community.

Axtell M.J., Westholm J.O., Lai E.C. (2011) Vive la difference: biogenesis and evolution of microRNAs in plants and animals. Genome Biol., **12:** 221.

Bashiardes S., Veile R., Helms C., Mardis E.R., Bowcock A.M., Lovett M. (2005) Direct genomic selection. Nature Methods, **1(2):** 63–69.

Bennett M.D., Smith J.B. (1976) Nuclear DNA amounts in angiosperms. Philos. Trans. R. Soc. Lond B. Biol. Sci., **274:** 227-274

Bushman F. (2003) RNA interference: applications in vertebrates. Mol. Ther., **7:** 9-10.

Chawade A., Sikora P., Brautigam M., Larsson M., Vivekanand V., Nakash M.A., Chen T., Olsson O. (2010) Development and characterization of an oat TILLING-population and identification of mutations in lignin and beta-glucan biosynthesis genes. *BMC Plant Biol.*, **10:** 86.

de Lange P., van Blokland R., Kooter J.M., Mol J.N. (1995) Suppression of flavonoid flower pigmentation genes in Petunia hybrida by the introduction of antisense and sense genes. Curr.Top. Microbiol. Immunol., **197:** 57-75.

Elbashir S.M., Lendeckel W., Tuschl T. (2001) RNA interference is mediated by 21- and 22-nucleotide RNAs. Genes Dev., **15:** 188-200.

Fire A., Xu S., Montgomery M.K., Kostas S.A., Driver S.E., Mello C.C. (1998) Potent and specific genetic interference by double-stranded RNA in *Caenorhabditis elegans*. Nature, **391:** 806-811.

Foubister V. (2003) New delivery vehicle for RNAi. Targets, **2:** 130-131.

Gadjieva R., Axelsson E., Olsson U., Vallon-Christersson J., Hansson M. (2004) Nonsense-mediated mRNA decay in barley mutants allows the cloning of mutated genes by a microarray approach. Plant Physiol. Biochem., **42:** 681-685.

Gilchrist E.J., O'Neil N.J., Rose A.M., Zetka M.C., Haughn G.W. (2006) TILLING is an effective reverse genetics technique for *Caenorhabditis elegans*. BMC Genomics, **7:** 262.

Gregory B.D., Yazaki J., Ecker J.R. (2008) Utilizing tiling microarrays for whole-genome analysis in plants. Plant J., **53:** 636-644.

Gregory R.I., Chendrimada T.P., Cooch N., Shiekhattar R. (2005) Human RISC couples microRNA biogenesis and posttranscriptional gene silencing. Cell, **123:** 631-640.

Hamilton A.J., Baulcombe D.C. (1999) A species of small antisense RNA in posttranscriptional gene silencing in plants. *Science*, **286:** 950-952.

Hannon G.J. (2002) RNA interference. Nature **418:** 244-251.

Henikoff S., Till B.J., Comai L. (2004) TILLING. Traditional mutagenesis meets functional genomics. Plant Physiol., **135:** 630-636.

Hoffmann G.R. (1980) Genetic effects of dimethyl sulfate, diethyl sulfate, and related compounds. Mutat. Res., **75:** 63-129.

Krysan P.J., Young J.C., Jester P.J., Monson S., Copenhaver G., Preuss D., Sussman M.R. (2002) Characterization of T-DNA insertion sites in *Arabidopsis thaliana* and the implications for saturation mutagenesis. OMICS, **6:** 163-174.

Krysan P.J., Young J.C., Sussman M.R. (1999) T-DNA as an insertional mutagen in *Arabidopsis*. Plant Cell, **11:** 2283-2290.

Lundqvist U. (2005) The Swedish Collection of Barley Mutants held at the Nordic Genebank. Barley Genetics Newsletter, **35:** 150-154.

McCallum C.M.C.L., Greene E.A., Henikoff S. (2000) Targeting induced local lesions in genomes (TILLING) for plant functional genomics. Plant Physiol., **123:** 439-442.

Morin R., Bainbridge M., Fejes A., Hirst M., Krzywinski M., Pugh T., McDonald H., Varhol R., Jones S., Marra M. (2008) Profiling the HeLa S3 transcriptome using randomly primed cDNA and massively parallel short-read sequencing. Biotechniques, **45:** 81-94.

Morohashi K., Xie Z., Grotewold E. (2009) Gene-specific and genome-wide ChIP approaches to study plant transcriptional networks. Meth. Mol. Biol., **553:** 3-12.

Ng S.B., Turner E.H., Robertson P.D., Flygare S.D., Bigham A.W., Lee C., Shaffer T., Wong M., Bhattacharjee A., Eichler E.E., Bamshad M., Nickerson D.A., Shendure J. et al. (2009)Targeted capture and massively parallel sequencing of 12 human exomes. Nature**, 461**: 272–276.

Nykanen A., Haley B., Zamore P.D. (2001) ATP requirements and small interfering RNA structure in the RNA interference pathway. Cell, **107:** 309-321.

Ozsolak F., Milos P.M. (2011) Single-molecule direct RNA sequencing without cDNA synthesis. *Wiley Interdiscip.* Rev. RNA. **2:** 565-570.

Paddison P.J., Caudy A.A., Hannon G.J. (2002) Stable suppression of gene expression by RNAi in mammalian cells. Proc. Natl. Acad. Sci. USA, **99:** 1443-1448.

Park P.J., Cao Y.A., Lee S.Y., Kim J.W., Chang M.S., Hart R., Choi S. (2004) Current issues for DNA microarrays: platform comparison, double linear amplification, and universal RNA reference. J. Biotechnol., **112:** 225-245.

Petersen D., Chandramouli G.V., Geoghegan J., Hilburn J., Paarlberg J., Kim C.H., Munroe D., Gangi L., Han J., Puri R., Staudt L., Weinstein J., Barrett J.C., Green J., Kawasaki E.S. (2005) Three microarray platforms: an analysis of their concordance in profiling gene expression. BMC Genomics, **6:** 63.

Plasterk R.H. (2002) RNA silencing: the genome's immune system. Science, **296:** 1263-1265.

Rose A.B., Beliakoff J.A. (2000) Intron-mediated enhancement of gene expression independent of unique intron sequences and splicing. Plant Physiol., **122:** 535-542.

Schena M., Shalon D., Davis R.W., Brown P.O. (1995) Quantitative monitoring of gene expression patterns with a complementary DNA microarray. Science, **270:** 467-470.

Shuey D.J., McCallus D.E., Giordano T. (2002) RNAi: gene-silencing in therapeutic intervention. Drug Discov. Today, **7:** 1040-1046.

Slade A.J., Fuerstenberg S.I., Loeffler D., Steine M.N., Facciotti D. A reverse genetic, nontransgenic approach to wheat crop improvement by TILLING. (2005) Nat. Biotechnol., **23:** 75-81.

Zakhrabekova S., Gough S.P., Braumann I., Muller A.H., Lundqvist J., Ahmann K., Dockter C., Matyszczak I., Kurowska M., Druka A., Waugh R., Graner A., Stein N., Steuernagel B., Lundqvist U., Hansson M. (2012) Induced mutations in circadian clock regulator *Mat-a* facilitated short-season adaptation

and range extension in cultivated barley. Proc. Natl. Acad. Sci. USA, **109:** 4326-4331.

Zakhrabekova S., Gough S.P., Lundqvist U., Hansson M. (2007) Comparing two microarray platforms for identifying mutated genes in barley (*Hordeum vulgare* L.). Plant Physiol. Biochem., **45:** 617-622.

Zakhrabekova S.K., Kannangra C.G., von Wettstein D., Hansson M. (2002) A microarray approach for identifying mutated genes. Plant Physiol. Biochem., **40:**189-197.

Zhang X., Yazaki J., Sundaresan A., Cokus S., Chan S.W., Chen H., Henderson I.R., Shinn P., Pellegrini M., Jacobsen S.E., Ecker J.R. (2006) Genome-wide high-resolution mapping and functional analysis of DNA methylation in *Arabidopsis*. Cell, **126:** 1189-1201.

Zhu J.Y., Sun Y., Wang Z.Y. (2012) Genome-wide identification of transcription factor-binding sites in plants using chromatin immunoprecipitation followed by microarray (ChIP-chip) or sequencing (ChIP-seq). Meth. Mol. Biol., **876:** 173-188.

Функциональный Геномикс, Прямые И Обратные Генетические Подходы Для Определения Важных QTL В Растениях

С.М. Захрабекова[1], С. Гоух[1], Л. Ланд[2], М. Хэнссон[1]

[1] *Лаборатория Карлсберга, Дания*
[2] *Отдел клеточной и молекулярной биологии, Университет Гетенборга, Швеция*

Основная цель функционального геномикса в том, чтобы понять, как генотип организма выражается в его фенотипа. Поскольку мутации могут препятствовать клеточным процессам, индуцированные мутации являются ключевым зондом для понимания функции гена.

Ключевые слова: *функциональный геномикс, прямая генетика, обратная генетика, QTL*

Funksional Genomiks Və Bitkilərdə Əhəmiyyətli QTL-ləri Müəyyən Etmək Üçün Birbaşa və Əksinə Genetik Yanaşmalar

Ş.M. Zəhrabəyova[1], S. Qoux[1], L.Lundh[2], M.Hansson[1]

[1]*Carlsberg laboratoriyası, Danimarka*
[2]*Göteborg Universiteti, Hüceyrə və molekulyar biologiya şöbəsi, İsveç*

Funksional genomiksin əsas vəzifəsi orqanizmin genotipinin öz fenotipində ifadə olunma formasının aydınlaşdırılmasıdır. Mutasiyalar hüceyrədəki proseslərə müdaxilə edə bildiyi üçün, genin funksiyasını bilmək üçün induksiya olunmuş mutasiyalar üçün açar rolunu oynayır.

Açar sözlər: *funksional genomiks, düzünə genetika, əksinə genetika, QTL*